

Introduction to the D.C. Cancer Registry

Purpose of the Registry

Population-based cancer registries are essential for assessing the extent of cancer burden in a specified geographic area. The District of Columbia Cancer Registry (DCCR) is a population-based cancer registry that collects incidence data on all cancer patients who reside in or who are diagnosed and/or treated for cancer in the District of Columbia. The goals of the DCCR are:

- to determine the incidence of cancer in the District of Columbia with respect to geographic, and demographic characteristics
- to monitor trends and patterns of cancer incidence over time
- to identify high risk populations
- to provide a database for and serve as a resource in conducting epidemiologic studies
- to provide data to assist public health officials, hospital administrators, and physicians to effectively plan services, prioritize health resource allocations and develop and measure prevention and intervention strategies.

History and Funding of the Registry

DCCR was established in 1985 following the introduction of home-rule for the self-governance of the District of Columbia. Prior to that date, the DCCR since 1963 was known as the Metropolitan Washington Cancer Registry with the mandate to gather the occurrence of cancer in the Washington metropolitan area covering Baltimore, Washington and Northern Virginia. The operations of the registry are mandated by Federal Statute code §6.4123 as amended in the Preventive Health Services Amendment

Act (D.C. Law 6-83 of 1985, and D.C. Law 8-157 of 1990). Additional funding has been awarded to DCCR from the Centers for Disease Control and Prevention (CDC) with a five-year federal grant aimed at enhancing the timely, complete, and accurate collection of cancer data, and for the computerization of data collection from reporting sources.

Collection of Data

Each District hospital, out-patient surgery center, and pathology laboratory is responsible for the complete ascertainment of all data on cancer diagnoses and treatments provided in its facility within six months of diagnosis. Sources for identifying eligible cases include:

- hospitals
- out-patient surgery centers
- private pathology laboratories
- free-standing radiation centers
- physicians whose cancer cases are not otherwise reported by the facilities listed above
- death certificate sources from State Center for Health Statistics (Vital Statistics Office)
- out-of-state cancer registries reporting a District resident receiving care in non-District health-care facilities.

When a cancer case is reported from more than one source the information is consolidated into one record. Reported cases contain the following data:

- patient demographics (including geographic place of residence at time of cancer diagnosis)
- description of cancer (including date of diagnosis, primary site, metastatic

sites, histology, extent of disease, etc)

- first course of treatment.

Primary site, behavior, grade, and histology are coded according to the “*International Classification of Diseases for Oncology, 2nd Edition*”. Stage of disease variables are coded using “*SEER’s Summary Staging Guide*” and “*AJCC Manual for Staging of Cancer, 4th Edition*”. All other variables are coded following the rules of the North American Association of Central Cancer Registries, the SEER program, and the American College of Surgeons.

Reportable Cases

All in-situ or malignant neoplasms are reportable to DCCR, except in-situ cancer of the cervix. The database includes all cases of carcinoma, sarcoma, melanoma, lymphoma, and leukemia, diagnosed by histology/cytology, radiology, laboratory testing, clinical observation, and autopsy.

Basal and squamous cell carcinomas of the skin are excluded except when occurring on a mucous membrane.

In conformance with guidelines provided by the North American Association of Central Cancer Registries, cervix in-situ cases, if reported to the DCCR, have been excluded from this publication.

Confidentiality of Data

The District of Columbia cancer reporting law ensures the protection of confidential data and restricts the release of identifying data. Persons with access to confidential data are required to sign a pledge of confidentiality and are subject to penalty if they, through negligence or willful misconduct, disclose confidential data.

Quality Assurance

To assure validity and reliability of the data presented in this report, DCCR has many mechanisms in place to check data for quality and completeness. We use EDITS software which have standard edits using algorithms that check the content of data fields against an encoded set of acceptable possible contents and flag the acceptability of coded data. Edits include field edits, inter-field edits, and inter-record edits. Edits check for unlikely sex/site, site/histology, or site/age combinations. In addition to computerized edits, each case is manually reviewed for errors.

Records are also routinely checked for duplicate entries. Duplicate case checking is performed both manually and electronically using various methodologies.

The District’s 1999 incidence data used for this report have been audited by the North American Association of Central Cancer Registries. NAACCR audits are performed on pre-determined sample of reporting facilities and cases in a state’s central cancer registry database for a given year. The procedure involves re-casefinding and re-abstracting studies on selected facilities and cases to determine the extent of completeness and accuracy of cancer data. In the NAACCR Audit of the District’s 1999 incidence data, all non-federal reporting hospitals were included in the sample of hospitals, and 10 percent of the reported cases were sampled for re-abstracting studies.

The 1999 incidence data has been submitted to NAACCR for review. This standard setting organization has given the 1999 DC incidence data a gold certificate performance. This indicates that the data has met the organization’s standards for

completeness of case ascertainment and of individual data items and accuracy of recorded information. This is the third consecutive year that DCCR has been awarded the “Gold” Certification.

Executive Summary

Data Presentation

The Report is compromised of eight parts. Part I is Foreword and Acknowledgments. Part II Background includes Introduction to the D.C. Cancer Registry, Executive Summary, and Technical Notes. Part III is “1999 Incidence and Mortality Age-adjusted against 2000 US Standard Population”, which is compromised of six sections. The first three sections are for incidence rates. Age-specific rates, age adjusted rates by race, gender, and ward of residence have been presented in these sections. Each of section contains the rates for 26 cancer site categories. The last three sections are for mortality rates of the same 26 cancer site categories in the same format. Part IV is “1999 Incidence and Mortality Age-adjusted against 1970 US Standard Population”. This part presents the same format of statistics as the Part III but using US 1970 standard population rather than US 2000 standard population. Part V uses Geographic Information System techniques to present 1999 Incidence and Mortality Maps by Ward for the 26 selected cancer site categories. Part VI is reference. Part VII is Appendix, which includes the distribution of both 2000 and 1970 US standard population, Estimation of DC total population and the population by ward. Part VIII provides the prevention guidelines for commonly diagnosed cancers from ACS.

This report has used the United States 2000 Standard Population for age adjustment. However, United States 1970 Standard Population has also been used for comparison purpose since all previous years’ rates have been adjusted using US 1970 standard population. Since the current age distribution of population in District of

Columbia is more close to the US 2000 standard population distribution compared to 1970 US standard population, the US 2000 population adjusted rates are very close to the observed rates. US 1970 standard rate weighs more on younger age groups, so its adjusted rates tend to be much lower than observed rates from the current populations with increased numbers in older age groups.

Population Description

The population of the District of Columbia in 1999 was estimated to be 521,064, made up of 244,005 males (46.8%), and 277,059 females (53.2%). Population estimates were obtained from the U.S. Bureau of the Census and National Cancer Institutes. The District of Columbia is comprised of 192 census tracts grouped into eight wards. The detailed population related information is given in the Appendix.

Descriptive Summary by Race and Gender

The data presented in this report cover those cases diagnosed among District of Columbia residents between January 1, 1999 and December 31, 1999. In this time frame, there were 3,013 cases of cancer diagnosed in the District, of which 2,067 were African Americans, 790 were Caucasians, 19 were Indians/Aleutians/Pacific Islanders. Rates are calculated for African Americans, Caucasians, and Indians/Aleutians/Pacific Islanders. Other races and subjects with unknown races are not included in any of the above race categories but are included in the calculation of rates for the total races. There were 1352 cancer deaths in 1999, including 675 males and 677 females. Among them, 1044 were African Americans, 291 were

Caucasians, and 13 were 19 were
Indians/Aleutians/Pacific Islanders.

Technical Notes

Age-specific Incidence Rates

Age-specific rates are calculated by dividing the number of cases for a given age-group by the total population of that age-group and are expressed as an average annual rate per 100,000 population by age group. Age-specific rates exclude the same types of cases that are excluded from age-adjusted incidence rates.

Let the age-distribution of deaths occurring in a given geographical area and year be denoted by d_i , and let the corresponding distribution of mid-year population at risk be denoted by p_i . By definition, therefore, the age-specific death rate for any age interval, m_i , can be expressed as:

$$m_i = \frac{d_i * 100,000}{p_i}$$

Eqn. (A)

where i denotes each of the N categories of age over which the deaths and population are distributed, and the computed rate is expressed per 100,000 population.

Direct Age-Adjusted Death Rate:

In order to permit the comparison of cancer mortality rates across populations whose age distributions might be different, it is recommended that age-adjusted rates using a standard population distribution be obtained. In this report the U.S. Standard Million Population, given in Appendix, Table 21, was employed to perform direct adjustments of the cancer mortality rates presented. The observed deaths and population were

grouped into the same eighteen (18) age intervals, 0-4, 5-9, 10-14, ..., 80-84, 85+ as the Standard Million Population. Thus, $i = 1, 2, \dots, 18$, and the age-adjusted incidence or mortality rate (R) can be expressed as:

$$R = \sum_{i=1}^{18} \frac{m_i p_i^s * 100,000}{P_T^s}$$

Eqn. (B)

where P_T^s is the sum of the standard population, p_i^s , over all 18 age intervals i.e.

$$P_T^s = \sum_{i=1}^{18} p_i^s$$

Eqn. (C)

The age-adjusted incidence and mortality rates published within this report were adjusted using the direct method and standardized to the age distribution of the U.S. 1970 population (see Appendix for the 1970 U.S. Standard Population). The rate represents the average number of new cases diagnosed annually per 100,000 persons. Age adjustment allows rates from one geographic area to be compared with rates from another geographic area that may have differences in age distributions. Any observed differences in age-adjusted incidence rates between populations are not due to differing age structures.

The computation of rates requires reliable estimates of the population at risk by five-year age groups, gender and race during the time period being studied. Population figures used in this report were obtained from U.S. Bureau of the Census estimates of

wards by age, gender and race, as released by the D.C. Office of Planning.

Confidence Interval of Age-Adjusted Death Rate:

The confidence interval of age-adjusted rates gives the estimated range of values within which the true population value lies with given probability. The confidence interval of the age-adjusted rate given in Equation (B) above may be obtained by employing the approximate expression given for the computation of the standard error of a crude mortality or incidence rate (Keyfitz, 1966):

$$C.I._{(R)} = R \pm c_{1-\alpha} S.E._{(R)}$$

Eqn. (D)

where: $S.E._{(R)} = R / [\text{events}]^{1/2}$

Eqn. (E)

and events denotes the number of deaths or incident cases upon which the calculation of the rate was based.

Risks and Associated Factors

The “risk and associated factors” subsections in Section I were developed from extracts from the following sources: 1995 Cancer Mortality Report of the District of Columbia, Cancer in Idaho 1996 Annual Report, Cancer Rates and Risks (NIH/NCI 4th Edition, 1996), Cancer Statistics for African Americans 1996, and the web-sites listed in the reference section.

Mean/Median/Mode

Measures of central tendency are helpful to describe a group of individual values in a simple and concise manner:

Mean: also known as the arithmetic average, is the sum of all observations divided by the number of observations.

Median: is the middle value when the observations are ranked in order from the smallest to the largest.

Mode: is the value which occurs most frequently in a group of observed values.

Cancer Case Definition

A “cancer case” is defined as a primary cancer site (where the cancer started), not a metastatic cancer site (where the cancer spread to). Since an individual can have more than one primary cancer site during their lifetime, the number of incident cancer cases are greater than the number of persons who are diagnosed with cancer.

Limitations to Data Interpretation and Comparison

Rates based on population estimates: District-wide and ward population figures are estimates, and errors in the estimates will also impact the rates.

Rates exclude missing/unknown demographic characteristics:

The numerators used in the calculation of age-specific and age-adjusted rates in this report exclude those cases for which the patient’s age, gender, or race are unknown or missing. As a result, occasionally tallies for “All Sites” may not equal the sum of the component categories. For example, due to missing gender, the sum of the number of cases for males and females may not sum up

to the total number of cases reported for “All Sexes”.

Rate comparisons: Age-adjusted incidence and age-specific rates based on a small number of cases (fewer than 10) may be unstable. In comparing rates among geographic areas (wards or other states), factors such as the absolute numbers of cases and differences in demographic characteristics should be considered. Interpretations without consideration of these factors may be misleading or inaccurate.

Standard Site Analyses Categories

To facilitate interpretation of data and comparisons across registries, DCCR uses standardized groupings of standard site analysis categories. These groupings are consistent with the National Cancer Institute’s SEER Program and are adopted by NAACCR. Most neoplasms are grouped by the organ where they occur. Neoplasms of the lymphatic, hematopoietic, as well as reticuloendothelial system are grouped by their histologies (leukemias, lymphomas, etc), and not by the anatomical site where they occurred (see Appendix for groupings of codes).